

An ECG Dataset Representing Real-World Signal Characteristics for Wearable Computers

Qingxue Zhang¹, Chakameh Zahed², Viswam Nathan⁴, Drew A. Hall³, Roozbeh Jafari⁴

¹University of Texas at Dallas, ²Texas Instruments, Inc.

³University of California, San Diego, ⁴Texas A&M University

qingxue.zhang@utdallas.edu, czahed@ti.com, viswamnathan@tamu.edu, drewhall@ucsd.edu, rjafari@tamu.edu

Abstract — We present an ECG dataset collected in real-world scenarios for wearable devices that includes over 260 recordings of 90-210 seconds that provide guidance for designers to evaluate signal acquisition circuit and system solutions. Several variations on the signal acquisition path are demonstrated, including various sources of interference (baseline wander, motion artifacts, and power line interference), signal path variations (electrode type, coupling method, and common-mode rejection method), and electrode placements (wrist and chest). Based on detailed analysis of signal characteristics under different scenarios, analog front-end (AFE) design recommendations are proposed.

Keywords—Wearable Computers; ECG Dataset; Motion Artifact; Baseline Wander; Powerline Interference; WEAR

I. INTRODUCTION

Wearable medical and health monitoring computers are playing an increasingly important role in our daily lives due to the convenience and continuous monitoring potential. Among all wearables, heart health monitors are gaining increased attention because heart disease has been the leading cause of death worldwide over the past decade [1]. Owing to the many meaningful temporal, spatial, spectral, and statistical features, electrocardiography (ECG) is regarded as the gold standard physiological signal for wearable heart health monitoring. In this paper, we focus on the wearable ECG-based heart health monitoring computers and provide a dataset that will guide design and development of such systems.

In most wearable applications, the signals are frequently impacted or corrupted by various environmental factors, such as baseline wander (BLW), motion artifacts (MA), and power line interference (PLI). Moreover, there are numerous variations in the acquisition system such as electrode type (wet vs. dry), coupling (AC vs. DC), and placement (wrist vs. chest) in addition to the common-mode rejection (CMR) method. Since all of these factors influencing the signal characteristics need to be considered when designing an analog front-end (AFE), a dataset incorporating a variety of wearable application scenarios provides valuable insight for designers to determine appropriate design specifications, such as the dynamic range, bit resolution and circuit activation policy.

Various useful ECG datasets exist today. PhysioBank, a large and open source archive of physiological signals, currently includes a large number of ECG datasets, most of which are clinical ECG data, such as *MIT-BIH Arrhythmia Dataset*, *Long-Term ST Dataset* and *Long-Term AF Dataset* [2]. Some datasets are for specialized conditions, like the

Abdominal and Direct Fetal ECG Dataset, and *MIT-BIH Noise Stress Test Dataset*. The latter is the only one that considers noisy ECG signals [2]. In addition, *American Heart Association (AHA) ECG Dataset* mainly provides clinical ECG data [3], and *Common Standards for Electrocardiography (CSE) Dataset* focuses on improving the ECG measurement standard and diagnostic methods [4].

Unfortunately, none of these datasets focus on wearable ECG applications; that is to say, no special attention has been paid to the unique conditions in wearable ECG computers, such as interference or electrode type, placement, and coupling. Although the MIT-BIH Noise Stress Test Dataset provides noisy recordings, they were synthesized by adding noise to clean ECG recordings, and do not necessarily represent *real-world* signals. In other datasets, the added noise is often simulated (e.g., 3 dB of white noise is added to ECG samples). Therefore, a dataset focusing on the wearable applications and representing a real-world experimental setup, along with true noise that we expect during real deployments, will be highly significant for analyzing the various signal characteristics and guiding design decisions. This dataset also provides the opportunity for algorithm developers to consider *real-world* data in algorithm design and evaluation.

The purpose of this article is to develop a novel wearable ECG dataset that provides *real-world* wearable ECG recordings, taking into account various sources of interference, signal path variations and electrode placements. Furthermore, after analyzing the signal characteristics, corresponding recommendations are proposed, aiming to help the designers to determine effective AFE solutions and appropriate design specifications. The Wearable ECG Activity Recordings (WEAR) dataset can be found at: <http://jafari.tamu.edu/wear>.

II. EXPERIMENTAL SETUP

We present global considerations in the ECG signal acquisition path, as shown in Fig. 1, and then present the use cases for the data collection and analysis, as shown in Table I.

A. Major Considerations in the Data Collection

1) *Activities*. 10 activities are considered, as shown in Table I (footnotes). Activity 1 is used to measure baseline wander, activities 2-8 are used to evaluate motion-induced artifacts, and activities 9 and 10 are used to measure power line interference at different distances to the power line. Motion-induced artifacts dominate the ECG noise during activities 2-8, hence baseline wandering was not considered during these activities.

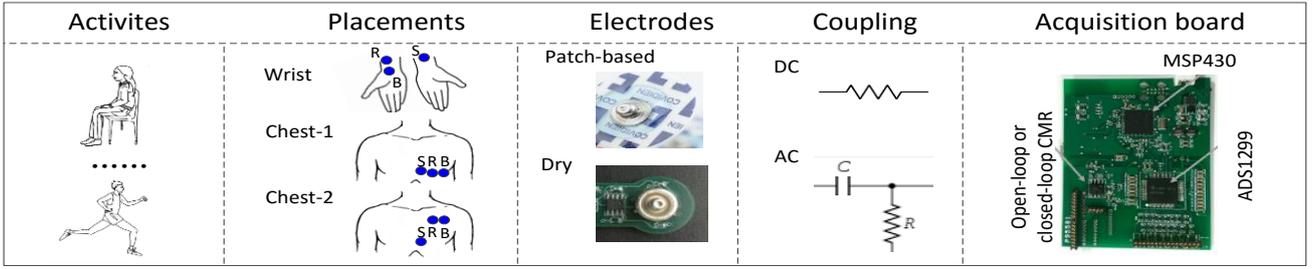


Figure 1. Global considerations in the ECG signal acquisition path: 1) 10 activities; 2) 3 electrode placements, including wrist/chest-1/chest-2 ECG (S/R/B represent signal/reference/bias electrodes); 3) 2 electrode types; 4) 2 coupling methods for the signal and reference electrodes; 5) the customized bio-potential acquisition board [7], providing two methods to drive the bias electrode for common-mode rejection (CMR): the open-loop and closed-loop CMR methods

2) *Electrode placements.* Single-lead ECG recordings, in particular from wrist and chest, are the most frequently used placements in wearable ECG applications due to their convenience. Besides wrist-based ECG, we also try two options for chest-based ECG, ‘chest-1’ and ‘chest-2’, with different distances between the signal electrode (S) and the reference/bias electrodes (R/B), to investigate two potential placements while maintaining acceptable signal quality. ‘chest-1’ placement, due to the closer proximity of electrodes, offers more convenient wearability.

3) *Electrode types.* Various electrode technologies have been reported and can be categorized into wet, dry, insulated, and non-contact electrodes for wearable applications [5]. The wet electrode usually provides better signal quality in terms of the noise and motion sensitivity. The other three types provide better comfort and utility in wearable applications. The non-contact electrode has an unresolved problem – susceptibility to motion-induced artifacts, and the insulated electrode is quite similar to the dry one in practice [5]. Therefore, we limit ourselves to dry and wet (patch-based) electrodes in this work.

4) *Coupling methods.* AC coupling removes DC offset as well as low frequency noise and interference, reducing the dynamic range requirement of the AFE. Here, the bias electrode is DC coupled, while the signal and reference electrodes are either both DC or both AC (1.6 Hz corner frequency) coupled to show the difference.

5) *Common-mode rejection (CMR) methods.* High CMR ratio (CMRR) is essential to reject the unwanted input signals common to both inputs of the differential amplifier. Two methods for CMR enhancement are considered in our data collection: open-loop and closed-loop CMR [6]. For the former, a 2.5 V common-mode voltage (bias voltage) directly drives the potential of the body through the bias electrode, while for the latter, an amplifier is used to sense and amplify the difference between the reference channel and this 2.5 V voltage and then drive the bias electrode, forming a closed-loop for CMR enhancement.

B. Data Acquisition Platform

A customized bio-potential platform, shown in Fig. 1, was used for data acquisition [7]. This includes an MSP430 microcontroller and a TI ADS1299 AFE with very low input-referred noise to preserve the original signal as much as possible. The sampling rate was set to 488 Hz [8]. A 24-bit ADC was used to minimize the quantization error and the gain

of the preamplifier was set to unity to reduce the possibility of saturation in the presence of large artifacts. Leveraging this setup, the influence of the AFE can be minimized to maintain the characteristics of the raw signal to the extent possible. AFE designers can set appropriate dynamic range targets based on recommendations given later. They can also lower the ADC resolution after evaluating the signal quality by adding quantization noise to the signals.

TABLE I. USE CASES FOR THREE CATEGORIES AND MEASURED MAXIMUM DYNAMIC RANGE (DR)

Categories	Connections			Place-ments	Activ-ities	Name	HB DR _{max}	Signal DR _{max}	Ratio
	Elec-trodes	CMR	Coup-ling						
Baseline wander (BLW)	Patch	CL	DC	Wrist	1	P_D_W	1.3	2.4	1.8
				Chest-1	1	P_D_C1	1.3	4.1	3.2
				Chest-2	1	P_D_C2	2.5	4.1	1.6
			AC	Wrist	1	P_A_W	1.1	1.2	1.1
				Chest-1	1	P_A_C1	1	1.1	1.1
				Chest-2	1	P_A_C2	2.3	2.3	1.0
	Dry	CL	DC	Wrist	1	D_D_W	1.4	33.1	23.6
			AC	Wrist	1	D_A_W	1.2	1.5	1.3
Motion artifact (MA)	Patch	CL	DC	Wrist	2-8	P_D_W	1.3	16.9	13.0
			AC	Wrist	2-8	P_A_W	1.1	7.5	6.8
	Dry		DC	Wrist	2-8	D_D_W	1.4	79.9	57.1
			AC	Wrist	2-8	D_A_W	1.2	14.4	12.0
Power line interference (PLI)	Patch	CL	DC	Wrist	9-10	P_CL_W	1.1	1.2	1.1
				Chest-1	9-10	P_CL_C1	1	1.1	1.1
				Chest-2	9-10	P_CL_C2	2.3	2.5	1.1
	AC		Wrist	9-10	D_CL_W	1.2	1.4	1.2	
			Wrist	9-10	P_OL_W	1.2	2.9	2.4	
			Chest-1	9-10	P_OL_C1	1	2.1	2.1	
	Dry	OL	Chest-2	9-10	P_OL_C2	2.3	3	1.3	
			Wrist	9-10	D_OL_W	1.3	2.8	2.2	

- 10 Activities: 1) sitting; 2) drinking coffee; 3) typing on the keyboard; 4) pressing and releasing the signal electrode; 5-6) walking at 1 and 3 mph; 7-8) running at 5 and 7 mph; 9-10) sitting, 0.6 and 6 feet to the power line.
- CL/OL CMR: closed-loop/open-loop common-mode rejection.
- HB DR_{max}: maximum peak-to-peak voltage of the clean heartbeat (HB) among 5 subjects. HB DR_{max} in category MA and PLI reuses that in category BLW.
- Signal DR_{max}: maximum peak-to-peak voltage of the signal after digital filtering among 5 subjects. Filtering methods:
 - BLW: 60 Hz notch filtering of the ECG signal with activity 1.
 - MA: 60 Hz notch filtering of the ECG signal with activity 8.
 - PLI: high pass filtering (1.6 Hz corner frequency) of the ECG signal with activity 9.
- Ratio: equals to ‘Signal DR_{max} / HB DR_{max}’.

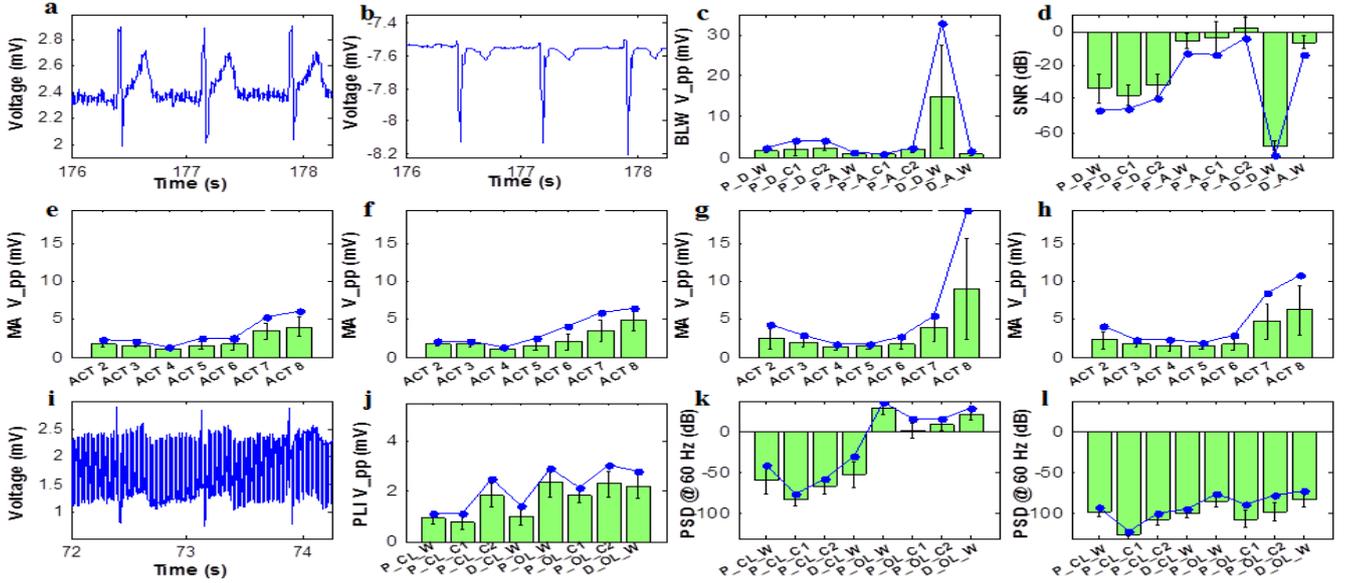


Figure 2. Data analysis for the baseline wander (BLW), motion artifact (MA) and power line interference (PLI) (a) – (d) for BLW: dry wrist (a) and chest-1 (b) ECG w/ DC coupling of subject 1; Max, mean and SD of ‘BLW V_pp’ (c); Min, mean and SD of SNR (d). Max: maximum; Min: minimum; SD: standard deviation; BLW V_pp: peak-to-peak voltage of signal contaminated by BLW after 60 Hz digital notch filtering; SNR: signal to noise ratio. (e) – (h) for MA: Max, mean and SD of ‘MA V_pp’ for different electrode types and coupling methods, patch + DC (e), patch + AC (f), dry + DC (g) and dry + AC (h). ACT: activity; MA V_pp: peak-to-peak voltage of signal contaminated by MA after digital high pass filtering (1.6 Hz corner frequency) and 60 Hz digital notch filtering. (i) – (l) for PLI: dry wrist ECG w/ OL CMR with activity 9 of subject 1 (i); Max, mean and SD of ‘PLI V_pp’ with activity 9 (j); Max, mean and SD of PSD at 60 Hz with activity 9 (k) and 10 (l). PLI V_pp: peak-to-peak voltage of signal contaminated by PLI after digital high pass filtering (1.6 Hz corner frequency); PSD: power spectrum density; CL/OL CMR: closed loop/open loop common-mode rejection.

C. Use Cases

The dataset currently includes 52 use cases for each of 5 subjects, 4 male and 1 female, all healthy individuals in the age group 20-35. There are a total of 260 recordings with duration of 90 to 210 seconds each with more recordings to be added in future. As shown in Table I, for baseline wander, different electrode types, coupling methods, and placements are considered, with a duration of 3.5 minutes for each recording. For motion artifacts, the emphasis is on various motion-related activities, with a duration of 1.5 minutes for each recording (at rest for the first and last 15s). For power line interference, the CMR methods, electrode placements and the distance to the power line are the major considerations, with a duration of 1.5 minutes for each recording. Therefore there are 94 minutes of data for each subject, and 140/210/120 minutes of data for baseline wander/motion artifacts/power line interference statistics. Each experiment is named as X_Y_Z where X ∈ {P, D} (P: patch-based electrode, D: dry electrode), Y ∈ {D, A, CL, OL} (D: DC coupling, A: AC coupling, CL: closed-loop CMR, OL: open-loop CMR), Z ∈ {W, C1, C2} (W: wrist, C1: chest-1, C2: chest-2). The measured dynamic range of the ECG signal without (HB DR_max) or with (Signal DR_max) interferences are shown in Table I and their ratio illustrates the enlargement of the dynamic range.

III. ANALYSIS

Based on all 260 ECG recordings, we provide both signal characteristics analysis (Fig. 2) and design guidelines for the low power AFE, emphasizing on the dynamic range (Table I).

Fig. 2 is based on the statistics of the 5 subjects and shows good consistency among subjects. The required bit resolution can be calculated based on the increase in dynamic range.

A. Baseline Wander (BLW)

Two examples of ECG signals are provided in Fig. 2a and 2b. Fig. 2c and 2d show the peak-to-peak voltage (BLW V_pp) and signal-to-noise ratio (SNR) of the ECG signal for 8 use cases (Table I). It is challenging to calculate SNR for ECG signals because we cannot precisely discriminate the signal from the noise and there is no concrete definition for the signal. However, prior investigations have recommended a number of methods for SNR calculation. One is described in [9]:

$$\text{SNR} = 10 * \log_{10} \frac{\text{avgpower}(\text{signal})}{\text{avgpower}(\text{blw})} \quad (1)$$

where $\text{avgpower}(\text{signal})$ and $\text{avgpower}(\text{blw})$ are the average power of the signal above 1.6 and below 1.6 Hz frequency separately. This definition is only applicable to baseline wander and 60 Hz interference, and will not work for motion artifacts due to the overlapping of ECG signal and motion artifact frequency bands at frequencies over 1.6 Hz.

1) *Placements.* The chest-1 ECG in Fig. 2b shows a good ECG morphology with several distinguishable waves, illustrating the feasibility of this convenient placement with three adjacent electrodes close to the heart, due to observing stronger electrical activity at the heart.

2) *Electrode types.* Compared with the patch-based electrode, the dry electrode with DC coupling (D_D_W)

corresponds to a higher BLW V_{pp} (33 mV) and a lower SNR of (-74 dB), while the one with the AC coupling (D_A_W) does not show significant difference.

3) *Coupling methods.* As shown in Fig. 2c and 2d, with the same placement and electrode type, the AC coupling shows a lower V_{pp} and a higher SNR over the DC coupling. The difference between the AC and DC coupling for the dry electrode is much more distinguishable.

4) *Dynamic range.* In Table I, with the patch-based electrode, the dynamic range for the DC coupling needs to be enlarged from 1.6 to 3.2 times with different placements, yet only 1.0 to 1.1 times for the AC coupling needed. With dry electrodes, 23.6 times enlargement is required for the DC coupling while the increase is 1.3 times for the AC coupling.

B. Motion Artifact (MA)

Fig. 2e to 2h illustrate the peak-to-peak voltage (MA V_{pp}) for 28 use cases (four different combinations of the electrode type and coupling method, and seven activities for each combination). The signal is processed by a 60 Hz digital notch filter, and a digital high pass filter (1.6 Hz corner frequency) for baseline wander removal in order to clearly demonstrate the motion-induced artifacts.

1) *Electrode types.* The patch-based electrode (Fig. 2e and 2f) corresponds to a smaller MA V_{pp} (motion artifact peak-to-peak voltage) compared with the dry electrode (Fig. 2g and 2h), benefitting from a better skin-electrode contact. A similar trend is observed in these 4 figures: with the increasing of motion intensity, more motion artifacts are induced, e.g., running (activity 7 and 8) corresponds to the highest MA V_{pp} , due to the drastic change of the contact impedance, EMG noise, and so on.

2) *Coupling methods.* The motion-induced artifacts are similar for both DC and AC coupling, when comparing Fig. 2e with 2f, and Fig. 2g with 2h, as the baseline wander was removed by either AC coupling or the digital filtering.

3) *Dynamic range.* In Table I, to provide real-world scenarios for circuit designers, the baseline wander is also considered, since when motion is introduced, the baseline wander may dramatically increase. Taking the running (activity 8) as an example, with patch-based electrode, the enlargement of the dynamic range is 13.0 times for the DC coupling and 6.8 for the AC coupling. Whereas with the dry electrode, they are 57.1 and 12.0 times, respectively.

C. Power Line Interference (PLI)

Fig. 2i to 2l illustrate an ECG waveform, peak-to-peak voltage (PLI V_{pp}) and power spectrum density (PSD) at 60 Hz for 16 cases of power line interference. The signal is processed by a digital high pass filter (1.6 Hz corner frequency) for baseline wander removal.

1) *Placements.* In Fig. 2k and 2l, the interference for the chest-1 and chest-2 placements is slightly lower than that of the wrist placement. For the dry and patch-based electrodes, no distinguishable difference is observed.

2) *Common-mode rejection (CMR) methods.* Compared with the open-loop CMR, the closed-loop CMR shows a

lower PSD at 60 Hz, *i.e.*, a better suppression of the power line interference (PLI), at both distances of 0.6 feet (Fig. 2k) and 6 feet (Fig. 2l) from the power line. The PLI V_{pp} (PLI peak-to-peak voltage) for 0.6 feet (Fig. 2j) also supports the same conclusion. When approaching closer to the power line (Fig. 2k), the performance difference of the two methods becomes more distinguishable. Fig. 2i gives an example of ECG with PLI, from which a huge interference can be observed. We are currently investigating a two-electrode setup (w/o the bias electrode) to evaluate the PLI suppression.

3) *Dynamic range.* The dynamic range enlargement is from 1.1 to 1.2 times for closed-loop CMR and from 1.3 to 2.4 times for open-loop CMR (Table I).

IV. CONCLUSIONS

In this paper, we introduced a novel wearable ECG dataset including 260 recordings, with 52 use cases for each of the five subjects, aiming to guide the design of the analog front-end for wearable computers. Eight use cases for baseline wander, 28 for motion artifact and 16 for power line interference are provided. These use cases consider electrode placements, electrode types, coupling methods and common-mode rejection methods. We proposed recommendations on the specifications of the AFE, especially the dynamic range, which needs to be enlarged by anywhere from 1.0 ~ 57.1 times according to specific scenarios. The dataset is made freely available to the scientific community: <http://jafari.tamu.edu/wear>.

REFERENCES

- [1] WHO, "The 10 leading causes of death in the world," <http://www.who.int/mediacentre/factsheets/fs310/en/>, 2014.
- [2] A. L. Goldberger, et al., "Physiobank, physiokit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.
- [3] AHA, "American heart association eeg database," https://www.ecri.org/components/Pages/AHA_ECG_DVD.aspx.
- [4] J. Willems, et al., "Common standards for quantitative electrocardiography: goals and main results. cse working party.," *Methods of Information in Medicine*, vol. 29, no. 4, pp. 263–271, 1990.
- [5] Y. M. Chi, T.-P. Jung, and G. Cauwenberghs, "Dry-contact and noncontact biopotential electrodes: methodological review," *Biomedical Engineering, IEEE Reviews in*, vol. 3, pp. 106–119, 2010.
- [6] V. Acharya, "Improving common-mode rejection using the right-leg drive amplifier," *Texas Instruments*, 2011.
- [7] V. Nathan, et al., "A 16-channel bluetooth enabled wearable EEG/EMG/ECG platform with dry-contact electrodes for brain computer interface," in *Proceedings of the 4th Conference on Wireless Health*, p. 17, ACM, 2013.
- [8] M. Raju, "Heart-rate and ekg monitor using the msp430fg439," *Texas Instruments*, 2005.
- [9] S. Agrawal and A. Gupta, "Projection operator based removal of baseline wander noise from eeg signals," in *Signals, Systems and Computers, 2013 Asilomar Conference on*, pp. 957–961, IEEE, 2013.